What is claimed is:

1. An information searching apparatus comprising:

an index information registering device that registers a word extracted from a document in association with the document as index information for document search;

a document searching device that searches a document corresponding to information relating to a requested search by referring to the index information registered by said index information registering device;

an unknown word extracting device that extracts an unknown word from the document being searched;

a document type determining device that determines a type of the document being searched; and

a registration permitting/inhibiting determining device that determines whether said index information registering device is permitted or inhibited to register the unknown word extracted by said unknown word extracting device as the index information, according to the type of the document determined by said document type determining device.

2. An information searching apparatus according to claim 1, wherein:

said document type determining device determines whether the document being searched is a character recognition processed document including character codes

obtained by a character recognition process, and

said registration permitting/inhibiting determining device is operable when the document being searched is the character recognition processed document, to inhibit said index information registering device from registering as the index information the unknown word extracted from the document being searched by said unknown word extracting device.

3. An information searching apparatus according to claim 1, wherein:

said document type determining device determines whether the document being searched is a character recognition processed document including character codes obtained by a character recognition process;

the information searching apparatus further comprises a permitting/inhibiting designating device operable when the document being searched is the character recognition processed document, to designate whether said index information registering device is permitted or inhibited to register as the index information the unknown word extracted by said unknown word extracting device from the document being searched; and

said registration permitting/inhibiting designating device is operable when the document being searched is the character recognition processed document, to determine whether said index information registering

device is permitted or inhibited to register as the index information the unknown word extracted by said unknown word extracting device, based on the designation by said permitting/inhibiting designating device.

5   4.   An information searching apparatus according to claim 1, wherein:

said document type determining device determines whether the document being searched is a character recognition processed document including character codes

10   obtained by a character recognition process; and

said registration permitting/inhibiting determining device is operable when the document being searched is not the character recognition processed document, to permit said index information registering device to

15   register as the index information the unknown word extracted from the document being searched by said unknown word extracting device.

5.   An information searching apparatus according to claim 1, wherein:

20   said document type determining device determines whether the document being searched is a character recognition processed document including character codes obtained by a character recognition process;

the information searching apparatus further

25   comprises a character correcting device operable when the document being searched is the character recognition processed document, to correct an error in the character

codes included in the document being searched; and

said unknown word extracting device extracts the unknown word from the document having the error corrected by said character correcting device.

5   6.   An information searching apparatus according to claim 1, wherein:

said document type determining device determines whether the document being searched is an image document;

10   the information searching apparatus further comprises a character recognition processing device operable when the document being searched is an image document, to obtain character codes from the image document by performing a character recognition process,

15 and a character correcting device that corrects an error in the character codes included in the document subjected to the character recognition process by said character recognition processing device; and

said unknown word extracting device extracts the

20 unknown word from the document being searched after correction of the document by said character correcting device.

7.   An information searching apparatus according to claim 5, wherein said character correcting device

25 refers to a character correction dictionary, in which characters having similar character patterns are stored in association with each other, and a word dictionary

for comparing character strings, and corrects the error by changing an associated character code included in the character recognition processed document to a character composing a character candidate string, out of character

5   candidate strings produced by adding at least one character pattern in the character correction dictionary to at least one character code included in the character recognition processed document, and matching a word included in the word dictionary.

10   8.   An information searching apparatus according to claim 1, wherein said document type determining device determines the type of the document being searched, based on at least one of a filename extension of the document being searched and attribute information

15   added in advance to the document being searched.

        9.   An information searching apparatus comprising:

        a word extracting device that extracts a word from a document;

        an index information registering device that

20   registers the word extracted from the document in association with the document as index information for document search;

        a document search device that searches a document corresponding to information relating to a requested

25   search by referring to the index information registered by said index information registering device; and

        a character correcting device that corrects an

error in the character codes included in a document

subjected to a character recognition process,

wherein said word extracting device extracts the

word from the document having the error corrected by

5   said character correcting device.

10.   An information searching apparatus according

to claim 9, wherein said character correction device

refers to a character correction dictionary, in which

characters having similar character patterns are stored

10   in association with each other, and a word dictionary

for comparing character strings, and corrects the error

by changing an associated character code included in the

character recognition processed document to a character

composing a character candidate string, out of character

15   candidate strings produced by adding at least one

character pattern in the character correction dictionary

to at least one character code included in the character

recognition processed document, and matching a word

included in the word dictionary.

20   11.   An information searching method comprising:

an index information registering step of

registering a word extracted from a document in

association with the document as index information for

document search;

25   a document search step of searching a document

corresponding to information relating to a requested

search by referring to the index information registered

in said index information registering step;

an unknown word extracting step of extracting an unknown word from the document being searched;

a document type determining step of determining a type of the document being searched; and

a registration permitting/inhibiting determining step of determining whether said index information registering step is permitted or inhibited to register the unknown word extracted in said unknown word extracting step as the index information, according to the type of the document determined in said document type determining step.

12. An information searching method comprising:

a word extracting step of extracting a word from a document;

an index information registering step of registering the word extracted from the document in association with the document as index information for document search;

a document search step of searching a document corresponding to information relating to a requested search by referring to the index information registered in said index information registering step; and

a character correcting step of correcting an error in the character codes included in a document subjected to a character recognition process,

wherein said word extracting step comprises

extracting the word from the document having the error corrected in said character correcting step.

13. A computer-executable information searching program comprising:

an index information registering module for registering a word extracted from a document in association with the document as index information for document search;

a document search module for searching a document corresponding to information relating to a requested search by referring to the index information registered by said index information registering module;

an unknown word extracting module for extracting an unknown word from the document being searched;

a document type determining module for determining a type of the document being searched; and

a registration permitting/inhibiting determining module for determining whether said index information registering module is permitted or inhibited to register the unknown word extracted by said unknown word extracting module as the index information, according to the type of the document determined by said document type determining module.

14. A computer-executable information searching program comprising:

a word extracting module for extracting a word from a document;

an index information registering module for registering the word extracted from the document in association with the document as index information for document search;

5   , a document search module for searching a document corresponding to information relating to a requested search by referring to the index information registered by said index information registering module; and

a character correcting module for correcting an

10   error in the character codes included in a document subjected to a character recognition process,

wherein said word extracting module comprises extracting the word from the document having the error corrected by said character correcting module.

15   15.   A computer-readable storage medium storing an information searching program comprising:

an index information registering module for registering a word extracted from a document in association with the document as index information for

20   document search;

a document search module for searching a document corresponding to information relating to a requested search by referring to the index information registered by said index information registering module;

25   an unknown word extracting module for extracting an unknown word from the document being searched;

a document type determining module for determining

a type of the document being searched; and

a registration permitting/inhibiting determining module for determining whether said index information registering module is permitted or inhibited to register

5 the unknown word extracted by said unknown word extracting module as the index information, according to the type of the document determined by said document type determining module.

16. A computer-readable storage medium storing an

10 information searching program comprising:

a word extracting module for extracting a word from a document;

an index information registering module for registering the word extracted from the document in

15 association with the document as index information for document search;

a document search module for searching a document corresponding to information relating to a requested search by referring to the index information registered

20 by said index information registering module; and

a character correcting module for correcting an error in the character codes included in a document subjected to a character recognition process,

wherein said word extracting module comprises

25 extracting the word from the document having the error corrected by said character correcting module.